DOCUMENT RESUME

ED 200 634 TM 810 221

AUTHOR Stallings, Jane A.
TITLE Measuring Program Implementation in Seven Follow Through Models.
PUB DATE Mar 81
NOTE 33p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Classroom Observation Techniques; *Comparative Analysis; *Compensatory Education; Evaluation Criteria; Grade 1; Grade 3; *Models; Primary Education; *Program Implementation
IDENTIFIERS *Classroom Observation Instruments; *Project Follow Through

ABSTRACT

A comprehensive observation study of classroom processes and instructional practices in Follow Through Planned Variation projects (FT) was completed in August of 1974. FT was evaluated primarily to determine whether educational innovators were able to implement their model programs of compensatory education in several diverse sites. Thirty six projects included in the sample represented seven sponsors of educational models, all geographic regions, urban and rural areas, several racial and ethnic groups, and a total of 136 grade 1 and 137 grade 3 classrooms. The Classroom Observation Instrument was employed to gather data of the following three types: (1) physical environment, materials, and seating arrangements; (2) grouping of children and teaching staff and recording of all classroom activities; and (3) interactions between adults and children. Each classroom was observed three days for five hours each day. The degree of implementation was defined as the extent to which a sponsor had been able to influence teachers in different sites to follow model specifications in their classrooms. Two criteria were used to judge implementation: (1) the similarity between the classrooms when compared on sponsor selected variables and (2) the difference between sponsor classrooms from non-FT classrooms. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED200634

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

MEASURING PROGRAM IMPLEMENTATION IN SEVEN FOLLOW THROUGH MODELS

Jane A. Stallings, Ph.D.
March 1981

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Stallings

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Stallings Enterprises, Inc.
409 Poppy Place
Mountain View, California 94043
(415)969-4264

TM 810 221

MEASURING PROGRAM IMPLEMENTATION IN SEVEN FOLLOW THROUGH MODELS

A comprehensive observation study of classroom processes and instructional practices in Follow Through Planned Variation projects was completed in August of 1974 by Stallings and Kaskowitz at Stanford Research Institute (SRI). The Follow Through Program was established by Congress in 1967 under the Office of Economic Opportunity when it became apparent that a program was needed in early grades of public school to reinforce and extend academic gains made by economically disadvantaged children enrolled in Head Start or similar pre-school programs. Project Follow Through was, and is, a planned variation research design; that is, the goal is to examine the differential effectiveness of programs based on divergent educational and developmental theories.

The primary aim of the Stallings and Kaskowitz (1974) evaluation of Follow Through planned variation was to determine whether educational innovators were able to implement their programs of compensatory education in several diverse sites. In the study reported here, the results of efforts by seven Follow Through sponsors* to deliver their models to 36 projects was examined. In this report, the degree of implementation is defined as the extent to which a sponsor has been able to influence teachers in different sites to follow model specifications in the classroom. Two criteria were used to judge implementation: (1) how similar are the classrooms to each other when compared on sponsor selected variables? and (2) how different are sponsor classrooms from Non-Follow Through classrooms? Implementation was evaluated by observing in the classrooms and coding observable events specified by the sponsors as critical to their model.

The study of implementation has too often been neglected in the evaluation of large-scale social reforms. A problem in educational innovations is that, in too many cases, the programs have not been implemented. Evaluation therefore could not yield information about either effectiveness or ineffectiveness of the innovation. Charters and Jones (1973, p. 6) call the collection of test data an "abrogation of professional responsibility" unless evaluators also measure, or at least describe differences between experimental and comparison programs (e.g., in the Follow Through program, to determine whether the behavior patterns of teachers and students are consistent with the planned innovation). The consequence of slighting the implementation factor is that "elaborately designed studies may end up as appraising non-events, with no one the wiser" (Charters and Jones, P. 5). Behind the Classroom Door (Goodlad et al., 1970) and "Planned Variation: From the Perspective of a Model Sponsor" (Weikart, 1973) provide information relevant to the implementation issue. Goodlad's study indicated that although teachers and principals of the sampled schools assumed that they had implemented educational innovations, findings from observations in the classroom revealed quite the reverse--that traditional education practices generally prevailed:

---

*Sponsors of educational models who were observed Spring 1973 were Far West Laboratory for Educational Research and Development, University of Arizona, Bank Street College of Education, University of Oregon, University of Kansas, High/Scope Educational Research Foundation, and Education Development Center.

"One conclusion stands out clearly: many of the changes we have believed to be taking place in schooling have not been getting into classrooms...there seems to be considerable discrepancy between teachers' perceptions of their own innovative behavior and the perceptions of observers" (Goodlad et al., 1970, pp. 97 and 98).

Weikart (et al., 1973) stated that Follow Through sponsors learned the hard way that there was "apparently a vast gulf between the smiles and nods of workshop sessions and actual classroom implementation of a model" (Weikart, p. 12). Despite "the enthusiasm of the summer workshops," Weikart found little change in the classrooms a few months later. These findings add confirmation to Mason's (1973) assertion that evaluation studies frequently find that the innovation "was not actually implemented in the manner specified by the developer."

Background

Project Follow Through began when researchers and other educational stakeholders were invited by the government to submit plans for establishing their various programs in public schools in order to test whether their individual approaches could improve the educational achievement of economically disadvantaged children. From the group that came forward, 22 were selected to implement their programs as Follow Through program sponsors. Eleven of the 22 sponsors had developed and tried their educational concepts in university settings; eight were affiliated with private research institutes and three were community developed programs.

The sponsors described their educational models to an audience that included representatives from school districts around the country at a conference in Kansas City, Kansas in 1968. Ultimately sponsor models were implemented in 154 Follow Through projects within 136 urban and rural communities located in all regions of the nation. The Follow Through sponsors then faced the challenge of program implementation--including guiding the behavior of teachers toward specified sponsor goals.*

Upon being chosen by a community, the sponsor had to develop a cooperative relationship with the local school officials, the parents, and the teachers. He also needed to specify the many aspects of the program, including the activities that should occur, the materials to be used, the use of classroom aides, and the approaches to be taken toward children. Field staffs had to be developed who could train classroom teachers. Some of the sponsor's training staff were from the local community and others traveled to the site from the sponsor's home location. Most sponsors had a two-week summer training session for teachers and aides, followed by weekly, semi-monthly, or monthly inservice training and supervision of the teachers.

It was recognized early in the Follow Through program that changing teacher behavior with a resultant change in children's behavior and test scores would take a long time. Thus, it was decided that the sponsors should have a five-year period in which to try to implement their models in the communities.

---

*Egbert (1973) provides a historical view of Follow Through.

In earlier evaluations of Follow Through, the major emphasis was to determine child outcome in terms of child performance. Yet it was clear that if such an effect was found, and if the effects were different from one model to another, we would not know what caused the differences. Therefore, we needed to know what was actually happening in the classrooms. In order to determine whether the sponsors were effective in getting teachers to practice their methods in the classroom, it was necessary to observe the classrooms systematically. We wanted to know whether a child's day in the classroom corresponded with the sponsor's educational prescriptions. It was evident that observations were needed to record what was occurring in Follow Through classrooms.

Sample for the Implementation Study

Of 22 Follow Through educational programs, seven were selected for observational study. Those chosen represent a wide spectrum of innovative educational theories. The range includes two models based upon positive reinforcement theory (from the University of Kansas and the University of Oregon), a model based primarily upon the cognitive developmental theory of Jean Piaget (High Scope Foundation), an open classroom model based upon the English Infant School theory (Education Development Center), and three other models drawn from Piaget, John Dewey, and the English Infant Schools (Far West Laboratory, University of Arizona, and the Bank Street College).

Four first-grade and four third-grade classrooms were observed in 36 cities and towns. This represented five projects for six Follow Through educational models and six projects for University of Arizona's model. One first-grade and one third-grade Non-Follow Through classroom were selected for comparisons at each project. The projects included in the sample represented all geographic regions, urban and rural areas, and several racial and ethnic groups. There was a total of 136 first-grade and 137 third-grade classrooms.

Observation of Classroom Process

The Classroom Observation Instrument developed at Stanford Research Institute was employed to gather data about classroom environment and processes. The instrument was initially developed in 1969 with the assistance of eight Follow Through sponsor representatives, with a goal of being flexible enough to record the salient features of a variety of programs.

Each classroom was observed three days for five hours each day. Two days focused upon teachers and aides, and one day focused upon a random selection of individual children.

The observation system is composed of three observation instruments: (a) identifies physical environment information, materials present, seating arrangements; (b) records the grouping of children and teaching staff and records all activities occurring in the classroom; (c) records interactions between adults and children (i.e., who speaks to whom, content of the conversation, and effect of the message. These interactions are recorded in five-minute segments, four times an hour).

Four dimensions of reliability have been examined in the main report of Stallings and Kaskowitz (1974). Classrooms were found to be acceptably stable on observed variables from one day to another. Observers were found to be 70 percent or more reliable on all codes used in forming variables. One code found to be less than 70 percent reliable was omitted from the study.

Assessment of Implementation

The study of implementation compared observed classroom processes with specific sponsor goals. No completely satisfactory way to approach this issue presented itself, since the Follow Through programs were not originally operationalized in terms of variables defined on SRI's classroom observation schedule. This problem was partially overcome by constant consultation with the program's sponsors during the creation of the variables. Also, each sponsor selected his own set of variables which reflected some of the critical components of his model. Implementation was judged by two criteria: (1) the uniformity of sponsor classroom scores on selected implementation variables and (2) how the sponsor classrooms differed from Non-Follow Through classrooms on these variables. The rationale for the second criterion was that Follow Through was intended to be an intervention program that offered alternatives to the conventional classroom. The primary focus on evaluating program implementation, then, should be on those essential components of each sponsor's model that differentiate that sponsor from the conventional classroom.

Admittedly, important dimensions of classroom implementation have not been recorded. For example, it was a goal for Far West Laboratory to have their teachers establish environments where a child could search for solutions to his problems in his own way and could risk, guess and make discoveries without serious negative psychological consequences; however, we could not record whether such environments had been established or whether children were solving problems. We could only record that children engaged in activities independently, worked with a variety of materials and asked questions. Nevertheless, past observation analyses of the Follow Through programs have been able to discriminate between programs and have indicated that the sponsors have been successful in training teachers and aides to observably perform as the sponsor desired (Stallings, Baker, Steinmetz, 1972; Stallings, 1973).

Method

Designation of the Sponsor Implementation Variables

From the several sections of the Classroom Observation Instrument, a list of variables intended to describe representative classroom elements was constructed by SRI staff for each sponsor's model. A Sponsor Variable Questionnaire was then compiled and sent to the individual sponsors for corrections, deletions, and additions.

What would be desirable for a closely controlled experiment would be an explicit statement by the sponsors of what proportion of the time critical variables would occur in an ideally implemented classroom. However, elements

in a classroom are not like those in a test tube. We have not yet learned to predict the amount of individualized attention, or feedback that a group of children need to meet specific goals. Each classroom group is made up of individuals, and individuals are most likely to need different rates of feedback or individualized instruction for maximum growth. Thus, even the most specific models such as University of Oregon and University of Kansas will alter the rate of reinforcement depending upon the need of the child. I the spring of 1973, the author (J. Stallings) visited ideal first and third grade classrooms specified by each sponsor. Although each teacher was an excellent example of the model, no sponsor was willing ultimately to suggest that all other first and third grade teachers should perform exactly like the one selected as the criterion. For instance, in East St. Louis in a University of Oregon first grade, the class spent the opening session enthusiastically discussing the movie Sounder which the children had seen on the previous day. The teacher was asking such questions as "How did you feel when the dog was shot?" and "Which of the people in the movie did you like best? Do you know anyone like those people?" A whole series of open-ended questions (coded as 2 on our observation instrument) were asked. The children responded and extended their responses. While this kind of discussion may occur within the University of Oregon model, the process is not part of the model specifications. Thus, it would seem unfair to use this classroom as a criterion against which to judge all other University of Oregon classrooms, even though the adults were clearly competent in using the skills and techniques prescribed in the University of Oregon model when teaching reading and math. We simply could not come up with a recommendation that would specify the rate of questions or reinforcement to be expected from even the most structured models.

In lieu of the ideal of exact specifications for each model, an appropriate and realistic alternative was to ask each sponsor to rate the variables as to (1) importance to the model and (2) expected frequency of occurrence relative to "conventional" classrooms, according to the following definitions:

(1) Importance to the Model--Rating Categories

Critical--This is an essential program component. Its absence in one of our classrooms would mean that the program was not implemented. It is emphasized in training the teaching staff.

Important--This program component is considered beneficial to children and the teaching staff is strongly encouraged to include this in the classroom; however, absence of this element would not indicate poor implementation.

Not Relevant--This is not a part of the curriculum we have developed. It is something with which we do not concern ourselves; the occurrence or lack of occurrence would not be important.

(2) Expected Frequency Relative to Conventional Classrooms--Rating Categories

More--This should occur more frequently in our classrooms than in conventional classrooms.

Same--This should occur about as often in conventional classrooms as in our classrooms.

Less--This should occur less frequently in our classrooms than in conventional classrooms.

Only variables that were considered critical and that should occur more often in the sponsors' classrooms were used as implementation variables. The amount of time allotted to each activity over a week's time was ranked by sponsors (as were the materials to be used in the activities and the preferred groupings of children and adults). Finally the sponsors were asked to add observation variables which they thought might provide more explicit information regarding the implementation of their model. The variables that were finally selected for each model from the Sponsor Variable Questionnaire are used in the sections that follow to assess model implementation. As can be seen on Table 1, there is considerable overlap in the variables chosen by the sponsors. However, it is the unique mix of variables that makes the models different from one another. As previously stated, all variables critical to sponsors' implementation have not been recorded. The variables selected represent the best descriptors the evaluators and sponsors could construct which could be observed reliably in sponsor classrooms.

Computing Implementation Scores

Two questions were addressed in this analysis:

- Are the classrooms of a given model similar to each other on selected variables?

- Are a sponsor's classrooms different from traditional classrooms on the implementation variables?

The first step in the assignment of implementation scores was to establish a standard of comparison for each implementation variable. Since Follow Through programs were intended to be innovative programs that offer alternatives to the conventional classroom, Non-Follow Through classrooms were used as the standard from which Follow Through classrooms should differ in specified ways. The standard was established separately for the first and third grades. Each Follow Through classroom was assigned an implementation score on each of the corresponding sponsor's implementation variables. The score is a number between 1 and 5 that represents the position of the Follow Through classroom mean or value relative to the distribution of Non-Follow Through classroom means or values. A nonparametric scaling technique was used rather than one that employs the mean and standard deviation of the Non-Follow Through classrooms because of the variety of distributions that were encountered for the Non-Follow Through classrooms. As an illustration, Figures 1 and 2 display histograms of first grade Non-Follow Through classrooms on two selected implementation variables as well as the mean and standard deviation. The distribution in Figure 1 has the familiar bell shape and the distribution in Figure 2 has a J shape. In Figure 2, one measurement is an outlier. Any parametric approach that may be

Table 1

LIST OF CRITICAL VARIABLES SELECTED BY SPONSORS

No.	Variables Description	Far West Labs	University of Arizona	Bank Street	University of Oregon	University of Kansas	High Scope	EDC
24	Child selection of seating and work groups	X	X	X			X	X
25	Games, toys, play equipment present	X	X	X			X	X
39	General equipment, materials present	X				X		X
65	Guessing games, table games, puzzles		X			X	X	
66	Numbers, math, arithmetic	X	X	X	X	X	X	X*
67	Reading, alphabet, language development	X	X	X	X	X	X	X*
70	Sewing, cooking, pounding		X				X	
71	Blocks, trucks						X	
74	Practical skills acquisition	X**		X				
83	[illegible]ety of activities, over one day	X	X	X			X	X
86	Teacher with one child	X	X	X			X	
87	Teacher with two children			X				
88	Teacher with small group		X	X	X	X	X	
92	Aide with one child	X		X				
94	Aide with small group		X		X	X	X	
114	One child independent	X	X	X				X
115	Two children independent							X
116	Small group of children independent	X		X				X
239	Math or science equipment/Academic Activities	X		X			X	X
240	Texts, workbooks/Academic Activities				X	X		
343	Child to adult, all verbal except response			X				
344	Individual child verbal interactions with adult	X	X	X	X	X	X	X
350	Child questions to adults	X	X	X			X	X
363	Child group response to adult academic commands/requests or direct questions				X			
372	Child presenting information to a group			X			X	
375	Adult instructs an individual child	X						X
376	Adult instructs a group				X			
390	Adult task-related comments to children						X	X
394	All adult acknowledgment to children	X		X	X	X	X	
398	All adult praise to children		X		X	X		
412	Adult feedback to child response to adult academic commands/requests, questions				X	X		
420	Adults attentive to a small group	X			X		X	
421	Adults attentive to individual children	X	X	X		X	X	
423	Positive behavior, adults to children	X	X	X				
435	Total academic verbal interactions				X			
438	Adult communication or attention focus, one child	X		X		X	X	X
440	Adult communication or attentive focus, small group				X		X	
444	Adult movement	X						
450	All child open-ended questions							X
451	Adult academic commands/requests and direct questions to children				X	X		
452	Adult open-ended questions to children	X	X	X			X	X
453	Adult response to child's question with a question						X	X
454	Child's extended response to questions		X	X				
456	All child task-related comments	X	X				X	
457	All adult positive corrective feedback	X			X	X		
460	All child positive affect	X	X					X
469	All adult reinforcement with tokens					X		
509	Child self-instruction, academic				X*			
510	Child self-instruction, objects			X			X	X
513	Child task persistence			X		X	X	
514	Two children working together, using concrete objects						X	
515	Small group working together, using concrete objects			X			X	
516	Social interaction among children	X		X				X
574	Child movement	X						X
599	Child self-instruction, nonacademic	X	X				X	
	Total number of Critical Variables	28** 27*	21	27	16** 17*	17	29	20** 22*

* Third grade only.

** First grade only.

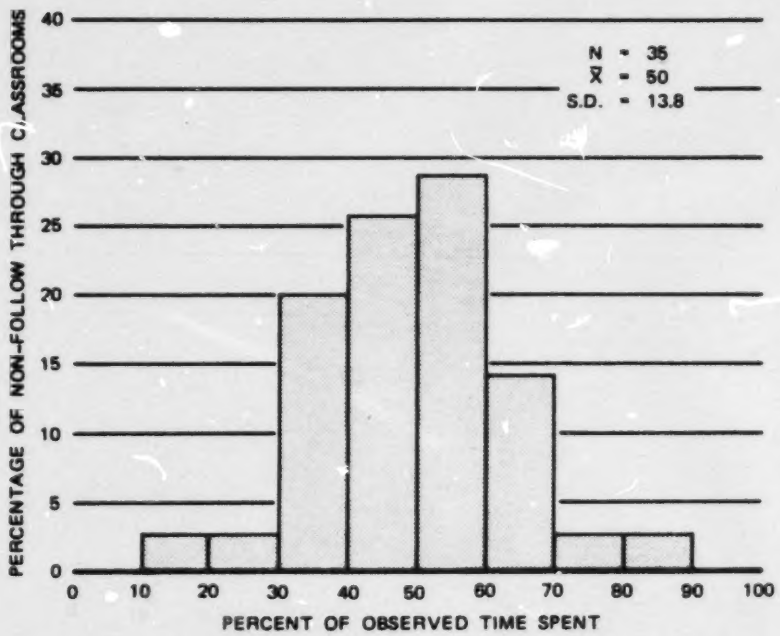

FIGURE 1 HISTOGRAM OF FIRST GRADE NON-FOLLOW THROUGH CLASSROOMS SHOWING PERCENT OF CHILD TIME SPENT IN READING, LANGUAGE (VAR. 67)

appropriate for one type of distribution may be inappropriate for another type. Also, the nonparametric procedure we elected to use tends to be less sensitive to outliers than a more conventional parametric procedure.

Several plans for establishing a standard means for comparison were considered and discussed with a panel of educational researchers. The previous SRI observation report (Stallings, 1973) had used quartile cutpoints in the distribution of Non-Follow Through classroom means to compute implementation scores for each Follow Through sponsor's classroom. More refined divisions of the Non-Follow Through distribution were considered such as using deciles; however, since there were only 35 Non-Follow Through first grade classrooms, only 3.5 Non-Follow Through classrooms would be represented in each decile. This suggestion was rejected on the basis that the precision of the estimation of the Non-Follow Through distribution did not warrant such a refined distinction in measuring implementation. As a reasonable compromise a quintile distribution of Non-Follow Through classroom means was selected to be the standard measurement.

Implementation scores for each sponsor were determined by rank ordering the Non-Follow Through classroom mean scores on each sponsor variable and dividing the distribution into five equal parts. There are 35 Non-Follow Through first grades. The seventh lowest score is the first quintile cutpoint; the fourteenth lowest score is the second quintile cutpoint; the twenty-first lowest score is the third quintile cutpoint; and the twenty-eighth lowest score is the fourth quintile cutpoint. Any Follow Through classrooms which had a score equal to or below the seventh score were in the first quintile. Any Follow Through classroom having a score above the twenty-eighth Non-Follow Through score was in the fifth quintile. Figure 3 shows the cutpoints for implementation scores for the variable "Games, toys, play equipment present" for the first grade Non-Follow Through classrooms. (See the Technical Note for a methodological discussion of the accuracy of quintile estimates.)

Quintiles:	1st	2nd	3rd	4th	5th
	7 Classroom Scores	7 Classroom Scores	7 Classroom Scores	7 Classroom Scores	7 Classroom Scores
	3.5 Cutpoint 1	4.2 Cutpoint 2	5.1 Cutpoint 3	6.3 Cutpoint 4	

FIGURE 3 NON-FOLLOW THROUGH FIRST GRADE CLASSROOMS--
GAMES, TOYS, PLAY EQUIPMENT PRESENT (VAR. 25)

Each sponsor's classrooms were assigned an implementation score according to the placement of the sponsor classroom means among the quintile cutpoints. If the classroom mean was greater than the fourth quintile cutpoint, an implementation score of 5 was assigned to the classroom; if the mean was less than or equal to the fourth quintile cutpoint, but greater than the third quintile cutpoint, a score of 4 was assigned. The quintile scores of 3, 2, and 1 were assigned in the same way. An implementation score of 5 indicates that the mean of the variable for the Follow Through classroom exceeded that of at least 28 classrooms, or 80%, of the Non-Follow Through classrooms. If a sponsor's classroom received an implementation score of 4 or 5, we could say that the classroom was on the upper end of the Non-Follow Through distribution; a score of 3 indicated that the classroom was in the mid-range of the Non-Follow Through distribution. The quintile cutpoints for each implementation variable are presented in Table 2.

There are some exceptions to the above specification of the implementation scores when several quintile cutpoints have a value of zero. This occurs when 14 (40%) or more of the Non-Follow Through classrooms have a mean of zero on a variable. (See Figure 4 for an example.) In such cases the rule was adopted that when the value for a Follow Through classroom was zero, the implementation score assigned to that classroom corresponds to the highest quintile cutpoint that was zero. For example, if the third quintile cutpoint was greater than zero and the second quintile cutpoint was equal to zero, then a Follow Through classroom with a value of zero would obtain an implementation score of 2 on that variable.

Quintiles:	1st	2nd	3rd	4th	5th
	7 Classroom Scores	7 Classroom Scores	7 Classroom Scores	7 Classroom Scores	7 Classroom Scores
	0.0 Cutpoint 1	0.0 Cutpoint 2	1.61 Cutpoint 3	3.70 Cutpoint 4	

FIGURE 4 NON-FOLLOW THROUGH FIRST GRADE CLASSROOMS--TEACHER WITH TWO CHILDREN (VAR. 87)

In several instances 28 (80%) or more of the Non-Follow Through classrooms, at a given grade level, had means of zero on six implementation variables (Variables 70, 74, 450, 469, 514, 515; see Table 2). In that case all quintile cutpoints were equal to zero; see Figure 5). Decision rules assigned an implementation score of 4 for these variables to a Follow Through classroom unless

Table 2

NON-FOLLOW THROUGH QUINTILES FOR SPONSOR IMPLEMENTATION VARIABLES

No.	Variables Names	Sponsor-Selected Variables* Far West** 1	Far West** 3	University of Arizona	Bank Street	University of Oregon** 1	University of Oregon** 3	University of Kansas	High/Scope	EDC	Non-Follow Through Quintiles: First Grade Minimum***	1st	2nd	3rd	4th	Maximum***	Third Grade Minimum***	1st	2nd	3rd	4th	Maximum***
24	Child selection of seating and work groups	X	X	X	X				X	X	0	0	0	1.0	1.0	2.0	0	0	0	1.0	2.0	4.0
25	Games, toys, play equipment present	X	X	X	X				X	X	1.0	3.0	4.0	5.0	6.0	9.0	1.0	2.0	3.0	4.0	5.0	7.0
39	General equipment, materials present	X	X					X		X	4.0	7.0	8.0	9.0	10.0	11.0	1.0	7.0	8.0	9.0	10.0	12.0
65	Guessing games, table games, puzzles			X				X	X		0	0	0	0	.51	7.82	0	0	0	.31	1.17	14.79
66	Numbers, math, arithmetic	X	X	X	X	X	X	X	X	X	0	11.23	16.97	19.32	23.26	32.7	9.71	16.01	20.12	23.85	26.56	38.70
67	Reading, alphabet, language development	X	X	X	X	X	X	X	X	X	19.50	38.82	46.36	54.31	58.61	84.20	37.56	38.61	45.84	52.18	55.41	73.52
70	Sewing, cooking, pounding			X					X		0	0	0	0	0	1.69	0	0	0	0	0	4.76
71	Blocks, trucks								X		0	0	0	0	.93	1.41	0	0	0	0	.45	1.80
74	Practical skills acquisition				X						0	0	0	0	0	.75	0	0	0	0	0	1.64
83	Wide variety of activities, over one day	X	X	X	X				X	X	2.00	4.00	4.67	5.33	6.25	7.33	2.67	4.33	4.67	5.00	5.67	8.67
86	Teacher with one child	X	X	X	X				X		0	0	1.85	4.55	11.86	61.11	0	0	1.64	4.41	11.76	58.82
87	Teacher with two children				X						0	0	0	1.61	3.70	10.00	0	0	0	0	4.35	24.59
88	Teacher with small group			X	X	X	X	X	X		0	6.90	11.67	21.67	30.98	88.06	0	0	9.43	14.63	22.41	42.86
92	Aide with one child	X	X		X						0	0	0	0	3.70	84.62	0	0	0	1.69	38.46	100.00
94	Aide with small group			X		X	X	X	X		0	0	0	15.38	42.86	80.65	0	0	0	0	22.58	100.00
114	One child independent	X	X	X	X					X	0	.25	1.05	1.69	3.66	6.41	0	.13	.57	1.06	2.21	10.63
115	Two children independent									X	0	0	1.17	2.26	4.20	8.23	0	0	.52	1.32	2.32	11.60
116	Small group of children independent	X	X		X					X	0	2.57	5.14	13.28	17.20	51.87	0	1.93	4.41	7.54	13.49	60.15
239	Math or science equipment/Academic Activities	X	X		X				X	X	0	0	0	5.56	10.53	44.00	0	0	0	0	5.26	26.53
240	Texts, workbooks/Academic Activities					X	X	X			18.18	41.86	57.69	73.68	85.55	98.25	7.32	40.54	62.75	78.57	87.27	100.00
343a	Child to adult, all verbal except response				X						0	1.17	1.85	2.53	4.16	7.25	.21	.82	2.41	3.41	4.95	10.32
344a	Individual child verbal interaction with with adult	X	X	X	X	X	X	X	X	X	12.40	20.55	23.75	27.31	29.89	42.47	7.83	20.20	23.49	28.67	33.26	48.10
350a	Child questions to adult	X	X	X	X				X	X	0	.38	.65	1.14	1.75	3.41	0	.52	.81	1.38	2.49	4.45
361a	Child group response to adult academic commands, requests, or direct questions					X	X				0	.89	1.89	2.43	3.22	5.00	.03	.45	1.00	1.24	2.10	3.14
372a	Child presenting information to a group				X				X		0	0	0	0	.14	3.35	0	0	0	0	.19	5.75
375a	Adult instructs an individual child	X	X							X	0	.50	.83	1.50	2.85	6.88	.13	.47	1.45	2.11	3.45	13.48
376a	Adult instructs a group					X	X				0	4.55	7.55	8.77	10.34	22.31	.13	2.40	5.32	7.70	10.55	20.35
390a	Adult task-related comments to children								X	X	0	.32	.57	1.05	1.90	11.21	0	.11	.59	1.17	1.95	6.57
394a	All adult acknowledgment to children	X	X		X	X	X	X	X		.50	1.44	2.29	2.95	3.65	7.88	.13	1.59	1.76	2.50	3.17	7.53
398a	All adult praise to children			X		X	X	X			.09	.49	.78	1.08	1.69	3.15	0	.17	.44	.64	1.03	2.47
412a	Adult feedback to child response to adult academic commands, requests, or questions					X	X	X			0	1.11	1.84	2.53	3.47	5.38	.05	.82	1.38	1.98	2.63	5.66
420a	Adult attentive to a small group	X	X			X	X		X		0	.05	.31	.84	1.32	6.56	0	.05	.10	.32	.66	3.53
421a	Adult attentive to one child	X	X	X	X			X	X		.03	1.05	2.56	3.60	5.10	16.62	0	2.26	3.14	4.90	7.47	10.87
423a	Positive behavior, adults to children	X	X	X	X						0	.05	.20	.39	.98	3.09	0	.03	.17	.33	.65	2.88
435c	Total academic verbal interactions					X	X				15.30	30.00	37.97	42.90	48.78	59.20	14.92	25.74	32.62	40.77	49.05	66.23
438c	Adult focus, one child	X	X		X			X	X	X	14.30	16.38	18.85	20.85	24.83	34.42	7.38	17.49	20.52	22.76	26.13	37.40
440a	Adult focus, small group					X	X		X		.03	1.02	2.46	3.91	7.82	15.90	0	.63	.92	2.17	3.97	10.18
444a	Adult movement	X	X								.05	.95	1.98	3.08	5.00	13.32	.13	1.73	2.69	3.89	5.18	12.10
450a	All child open-ended questions									X	0	0	0	0	0	.05	0	0	0	0	0	.10
451a	Adult academic commands, requests, and direct questions to children					X	X	X			3.10	5.18	6.84	7.78	9.02	13.31	1.23	3.53	5.13	7.66	9.21	12.80
452a	Adult open-ended questions to children	X	X	X	X				X	X	0	.10	.17	.23	.41	1.75	0	.03	.16	.45	.58	1.41
453a	Adult response to child's question with a question								X	X	0	0	0	.05	.13	.70	0	0	0	.03	.14	.71
454a	Child's extended response to questions			X	X						0	.06	.13	.47	.88	2.62	0	.05	.20	.51	.80	2.72
456a	All child task-related comments	X	X	X					X		0	.55	1.45	1.82	3.00	10.24	0	.49	1.08	1.90	3.21	9.88
457a	All adult positive corrective feedback	X	X			X	X	X			.58	1.21	1.67	2.20	3.00	4.57	.13	.77	1.49	2.00	2.55	4.55
460a	All child positive affect	X	X	X						X	0	.05	.15	.41	1.05	4.37	0	.06	.13	.30	.65	6.57
469a	All adult reinforcement with tokens							X			0	0	0	0	0	.15	0	0	0	0	0	.13
509c	Child self-instruction, academic					X	X				0	6.25	13.05	16.58	23.60	36.45	6.30	12.15	16.28	18.70	24.65	38.81
510c	Child self-instruction, objects				X				X	X	0	0	0	.25	2.95	19.94	0	0	0	0	4.90	23.50
513c	Child task persistence				X			X	X		0	1.75	2.05	2.90	4.20	6.80	0	1.80	2.60	3.24	5.90	6.56
514c	Two children working together, using concrete objects								X		0	0	0	0	0	0	0	0	0	0	0	.50
515c	Small group working together, using concrete objects				X				X		0	0	0	0	0	0	0	0	0	0	0	0
516c	Social interaction among children	X	X		X					X	.20	1.00	1.35	1.85	4.35	9.91	0	.90	2.26	3.16	4.45	11.25
574c	Child movement	X	X							X	.20	.80	1.20	7.32	2.94	4.80	0	.55	1.15	1.78	2.75	5.65
599c	Child self-instruction, nonacademic	X	X	X					X		0	0	.50	.89	4.60	10.89	0	0	.28	2.25	6.20	20.15

*An X indicates a variable selected by a sponsor.

**Separate variables for first and third grades are shown for Far West Lab and University of Oregon.

***Minimum and maximum represents the lowest and highest scores for Non-Follow Through.

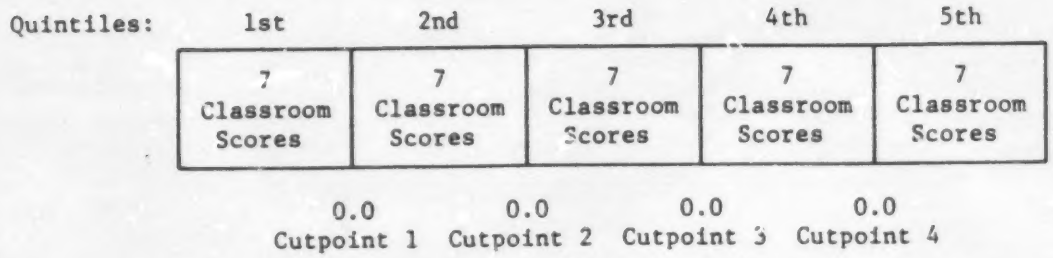

FIGURE 5 NON-FOLLOW THROUGH FIRST GRADE CLASSROOMS--SEWING, COOKING, POUNDING (VAR. 70)

the classroom mean was above zero. Any classroom that had a mean above zero on these variables would receive an implementation score of 5. In any case the recorded implementation score of 4 could be misleading and scores should be interpreted with caution. It must be remembered that even though some variables are infrequent, it is important to note whether they occur in sponsored classrooms and not in the Non-Follow Through classrooms. (See Table 2 to identify those variables where two or more quintile cutpoints were the same.)

A total implementation score for a classroom was computed by summing the implementation scores across the corresponding sponsor's implementation variables and then dividing by the highest possible sum. The resulting proportion was then multiplied by 100 so that the total implementation score is expressed in terms of a percentage of the total possible. For example, if a hypothetical sponsor's classroom were being rated on four variables, the highest possible sum of implementation scores for a classroom would be 4 x 5 = 20. If a classroom had implementation scores of 3, 3, 4, and 5 on the individual implementation variables, then the total implementation score for the classroom would be (15/20)100 = 75%. The reader needs to understand that there is no zero point when computing implementation scores. If all classrooms received the lowest implementation score of 1 on every single implementation variable, their overall implementation score would be 20%, not zero. If they received scores of 5, their overall implementation score would be 100%. Thus, the actual range is from 20 to 100 and the midpoint is a score of 60.

In order to compare the sponsor programs with Non-Follow Through classrooms, a total implementation score was also computed for each Non-Follow Through classroom on each sponsor's set of implementation variables. The mean and standard deviation of the Non-Follow Through pooled classrooms are reported for each sponsor. These statistics serve three purposes:

- The Non-Follow Through mean serves as a reference point for an implementation score. If all the quintile cutpoints had been distinguishable, then the mean total implementation score for the Non-Follow Through classrooms would be 60 (an average score of 3 on each variable out of 5 possible on all variables). Since

there are a number of variables where some or all Non-Follow Through quintiles are zero, it was necessary to compute a separate mean for each set of sponsor-implemented variables.

- The standard deviation was used as a scaling factor to compare a sponsor's total implementation score relative to the Non-Follow Through score. A t-test was used to test whether the mean total implementation score for Follow Through was significantly greater than the mean for Non-Follow Through.

An analysis of variance was run separately for each sponsor and grade level to test whether the sites differed on total implementation scores. This test indicates whether the variability in implementation scores among sites is large relative to the within-site variability among classrooms at a site.

Results

The report of implementation findings was provided for each sponsor in separate sections in the Follow Through Classroom Observation Evaluation, Stallings and Kaskowitz, 1974. For this paper data from only the Direct Instruction model will be analyzed to provide an example of how program implementation was evaluated.

Direct Instruction Model--University of Oregon

Description of the Model

The Direct Instruction model developed by Z. Engleman and W. Becker is a highly structured academic program based on the premise that with proper instruction and consistent reinforcement, any child can master the skills necessary to bring him up to the achievement level of national norms for his age group.

The model uses programmed reading, math, language, and science materials which they have developed. Highly specific methods are used to teach concepts and skills required for mastering sequenced tasks oriented toward an increasing level of competence. Desired behaviors are systematically reinforced by praise and enthusiastic acknowledgment. Teachers and students share the pleasure of each other's achievements. Unproductive or antisocial behavior is ignored or stopped by a short reprimand. Ignoring is intended to withdraw from the behavior the attention that ofter reinforces it. When this is true, ignoring will eventually extinguish the behavior.

The classroom is usually staffed with two or three adults (regular teacher and one or two full-time aides recruited from the Follow Through parent community) for every 25 to 30 children. Each adult has been carefully trained. In most cases the adults teach a single subject. Working very closely with a small or large group of children, each teacher and aide uses the programmed materials in combination with predefined teaching strategies. A task is presented by the teacher. A demonstration is usually given and performance of the task by the children is tested. When the children respond she provides immediate feedback. She proceeds only when each child is successful with a given instructional unit. When the children have completed their lessons with one adult, they proceed to their next subject with a different adult.

Description of the Sites

The University of Oregon project sites were East St. Louis, Illinois; New York City; Racine, Wisconsin; Tupelo, Mississippi; and Providence, Rhode Island. Demographic and other information describing the five communities in which the University of Oregon projects were located (taken from U.S. Census Bureau reports and Follow Through sources) is shown in Table 3.

Tupelo was the only one of the University of Oregon project sites in this study that was located outside of a metropolitan area; it was also the only University of Oregon site that had a higher percent of adults in the community who had completed high school than the Follow Through sample average.

Racine's $10,526 median family income was higher than the other University of Oregon sites and was, in fact, higher than any other site in the observation sample; its 6.6% of families below poverty level* was among the lowest over all sites.

The East St. Louis median family income of $6,654 was among the lowest in the sample, and, similarly, the 28.5% of families below the poverty level was one of the highest percentages in the sample. The nonwhite population of East St. Louis was 69.3%, a considerably higher percentage than all but one of the project sites in the Follow Through sample.

New York City was the only large city in the University of Oregon sample. Its $9,682 median family income was above the sample average and its 11.5% of below poverty level families was about average compared to the other cities. New York City's nonwhite population of 23.4% was slightly above the average for all Follow Through sites. It should be noted that this figure reflects all of New York City: The University of Oregon classrooms in New York were 95% black.

Although Providence was a medium-sized city in a metropolitan setting, demographic data for Providence look surprisingly like those for Tupelo with respect to median family income and percent of families below poverty level. The nonwhite populations of the two sites differed by 7.6%, yet both were below the average for all project sites.

Sponsor Implementation Variables

Tables of all the critical variables selected to assess implementation of the University of Oregon model can be found in Table 1. As described in the methodology section of this report, the final selection of variables was made on the basis of sponsor ratings reported in the Sponsor Variable Questionnaire. The variables selected for the University of Oregon model are of course limited descriptors of the program. A number of processes and procedures important to the program are not assessed in this study of implementation.

Sixteen critical variables were selected for first grade and 17 for third grade by University of Oregon as those on which they would expect their classrooms to be different from conventional classrooms. An implementation score of 3 would indicate the classroom was in the middle of the scale for the conventional classrooms; an implementation score of 5 would mean that the model

*The U.S. Census Bureau uses the U.S. Office of Economic Opportunity's poverty index guidelines to establish poverty level.

Table 3

SITE AND DEMOGRAPHIC DATA FOR UNIVERSITY OF OREGON SITES

A. Geographic Data

Site Code	Site Location	Region	Metropolitan Status	Population (1970 census)
0703	E. St. Louis, Ill.	West North Central	Medium city	69,996
0707	New York City, P.S. 137K	Middle Atlantic	Big city	7,894,798
0708	Racine, Wis.	East North Central	Medium city	95,162
0711	Tupelo, Miss.	East South Central	Small city*	20,471
0719	Providence, R.I.	North East	Medium city	179,213

B. Demographic Data for Total Population

Site Name	Median Family Income	Percent of Families Below Poverty Level	Percent of Population Nonwhite	Percent of Adults Over 25 Who Finished High School
E. St. Louis	$ 6,654	28.5%	69.3%	29.4%
NYC P.S. 137K	9,682	11.5	23.4	46.9
Racine	10,526	6.6	11.0	49.5
Tupelo	8,436	12.5	17.6	58.9
Providence	8,430	13.3	10.0	40.6
Average for all sites	$ 8,631	13.9%	20.7%	49.7%

C. Characteristics of the Follow Through Evaluation Sample at the Site for First Grade**

Site Name	Average with Preschool	Average Percent with First Language Other than English	Average Baseline Score
E. St. Louis	47%	0%	--
NYC P.S. 137K	43	2	29
Racine	75	9	--
Tupelo	63	0	--
Providence	97	0	26

*Not within Standard Metropolitan Statistical Area (SMSA).

**Taken from the Follow Through Roster; represents the Follow Through evaluation sample.

Sources: U.S. Bureau of the Census, 1970; SRI

classrooms were in the uppermost range of conventional classrooms. A selection of variables designated as critical by the sponsor are discussed in the text.

Implementation Findings

Figure 6 and Table 4 present the total implementation scores for each classroom by grade level. These total scores were computed by adding each quintile score of the University of Oregon's implementation variables and dividing by the total maximum score. Thus, an average score on all variables was computed. The scores are presented as percentages. Total scores for Non-Follow Through classrooms were also computed for the University of Oregon implementation variables. Means and standard deviations are presented in Table 4 along with the results of one-tailed t-tests, comparing Non-Follow Through classrooms with classrooms using the University of Oregon model. Table 4 presents the analysis of variance among sites.

As the t-test on Table 4 indicates, overall, the University of Oregon's classrooms were significantly different from the Non-Follow Through classrooms. The histograms on Figure 6 show that only one Oregon first grade had a score lower than 70 while 25 Non-Follow Through classrooms had scores lower than 70. In the third grade only one Oregon classroom scored below the mean of Non-Follow Through.

The analysis of variance indicates that the among-site variance was greater than the within-site variance in the first grade. In the third grade the variance was as great within sites as among sites. The difference in New York was particularly great where one third grade had a low score of 57.6 and another had a high score of 81.2. A possible explanation may be that children in University of Oregon's New York third grades had fewer months in Follow Through and the attrition rate was greater. The standard deviation for third grades in St. Louis and Racine was also high. The standard deviation for the first grades at these same sites was considerably less. This difference between the grades might be explained by the fact that in the spring of the year in third grade when observations were conducted teachers were beginning to prepare the children for the fourth grade Non-Follow Through classrooms and they might not have been adhering so strictly to University of Oregon's stated program.

Scores for each critical implementation variable for each University of Oregon site and classroom are presented in Appendix A in tabular form. The tables are prepared so that the number of teachers receiving a particular implementation score is entered in a column. The number of teachers from each site is entered in the rows. The total number of teachers receiving a particular implementation score, as well as the percentage computed, is entered on the bottom rows. Only a few critical variables will be discussed in this text.

The University of Oregon program emphasized the development of basic skills in reading and computation. Findings presented in Table 5 indicate that 90% of the first grade classrooms in the University of Oregon program often engaged in arithmetic activities and were on the upper end of the scale when compared with traditional classrooms. While three third grade classrooms in East St. Louis and four classrooms in Tupelo had the highest implementation

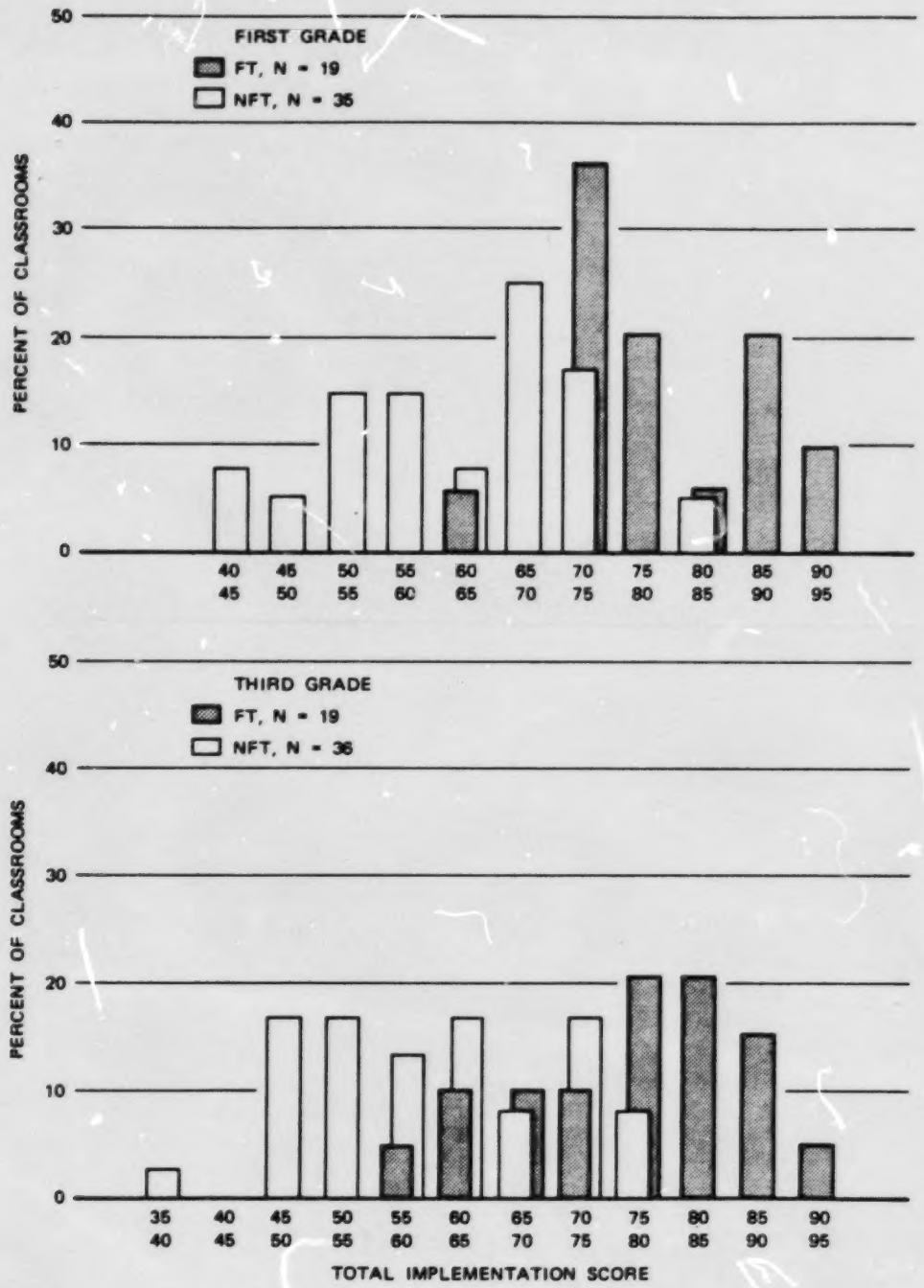

FIGURE 6 HISTOGRAM SHOWING IMPLEMENTATION SCORES FOR UNIVERSITY OF OREGON

Table 4

TOTAL IMPLEMENTATION SCORES FOR CLASSROOMS BY SITE--UNIVERSITY OF OREGON

First Grade

Sites	Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
E. St. Louis (EK)	76.2%	62.5%	76.2%	75.0%	72.5%	6.7
NYC P.S. 137K (EK)	88.7	90.0	91.2		90.0	1.3
Racine (EK)	72.5	72.5	71.2	71.2	71.9	.7
Tupelo (E1)	80.0	86.2	87.5	87.5	85.3	3.6
Providence (EK)	72.5	77.5	72.5	73.7	74.1	2.4
Sponsor Scores (N=19):					78.2%	8.1
NFT Scores (N=35):					61.0	10.7
					t = 6.11	
					p < .001	
					f = 17.61	
					p < .001	

Third Grade

Sites	Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
E. St. Louis (EK)	76.5%	62.4%	78.8%	87.1%	76.2%	10.3
NYC P.S. 137K (EK)	68.2	81.2	57.6		69.0	11.8
Racine (EK)	71.8	62.4	84.7	85.9	76.2	11.2
Tupelo (E1)	87.1	80.0	90.6	74.1	82.9	7.3
Providence (EK)	75.0	82.4	78.8	69.4	76.4	5.5
Sponsor Scores (N=19):					76.5%	9.3
NFT Scores (N=36):					60.4	10.5
					t = 5.62	
					p < .001	
					f = .91	
					p < NS	

Table 5

NUMBERS, MATH. ARITHMETIC (Variable 66)--UNIVERSITY OF OREGON

Sites	First Grade Classrooms with Implementation Scores of					Third Grade Classrooms with Implementation Scores of				
	1	2	3	4	5	1	2	3	4	5
E. St. Louis, Ill.					4			1		3
NYC P.S. 137K					3		1	2		
Racine, Wis.				2	2			2	2	
Tupelo, Miss.				1	3					4
Providence, R.I.		2			2		2	1		1
Total classrooms		2		3	14		3	6	2	8
Percent of classrooms		11%		16%	74%		16%	32%	11%	42%

score, all but one third grade classroom in other sites were scattered across scores 2, 3, and 4. It must be noted that Tupelo was a site where children entered school in the first grade while at all other sites the children entered school in kindergarten.

As shown in Table 6, 63% of University of Oregon's third grade classrooms received the highest implementation score of 5 in reading and language development. The first grade implementation scores were scattered, with 47% obtaining a score of 3. This indicates that while first grades differed more from the traditional classrooms in how often children engaged in math, the third grades differed more from the traditional classrooms in how often the children engaged in reading.

For instructional purposes, University of Oregon children were placed in small groups. Findings presented in Tables 7 and 8 indicate that both teachers and aides in University of Oregon first and third grade classrooms worked more often with small groups of children than did teachers in comparison classrooms.

In general, in the University of Oregon program, after adults gave instructions to the small group, they then asked a question regarding the information given. The children responded, and the adult provided immediate feedback. The findings for the variable representing this instructional pattern

Table 6

READING, ALPHABET, LANGUAGE DEVELOPMENT (Variable 67)--UNIVERSITY OF OREGON

Sites	First Grade Classrooms with Implementation Scores of					Third Grade Classrooms with Implementation Scores of				
	1	2	3	4	5	1	2	3	4	5
E. St. Louis, Ill.		1	1	2			1	1	1	1
NYC P.S. 137K				1	2	1			2	
Racine, Wis.			4							4
Tupelo, Miss.		1			3				1	3
Providence, R.I.			4							4
Total classrooms		2	9	3	5	1	1	1	4	12
Percent of classrooms		11%	47%	16%	26%	5%	5%	5%	21%	63%

Table 7

TEACHER WITH SMALL GROUP (Variable 88)--UNIVERSITY OF OREGON

Sites	First Grade Classrooms with Implementation Scores of					Third Grade Classrooms with Implementation Scores of				
	1	2	3	4	5	1	2	3	4	5
E. St. Louis, Ill.	1				3		1		1	2
NYC P.S. 137K				1	2	1		1		1
Racine, Wis.					4	2			1	1
Tupelo, Miss.					4					4
Providence, R.I.					4					4
Total classrooms	1			1	17	3	1	1	2	12
Percent of classrooms	5%			5%	89%	16%	5%	5%	11%	63%

Table 8

AIDE WITH SMALL GROUP (Variable 94)--UNIVERSITY OF OREGON

Sites	First Grade Classrooms with Implementation Scores of					Third Grade Classrooms with Implementation Scores of				
	1*	2	3	4	5	1*	2*	3	4	5
E. St. Louis, Ill.			1	1	2					4
NYC P.S. 137K				1	2				2	1
Racine, Wis.					4			1	2	1
Tupelo, Miss.					4					4
Providence, R.I.				1	3				1	3
Total classrooms			1	3	15			1	5	13
Percent of classrooms			5%	16%	79%			5%	26%	68

*Of the Non-Follow Through classrooms, 20% in the first grade (the lowest quintile) had a mean of zero and 40% in the third grade (th two lowest quintiles) had mean of zero. No University of Oregon s had a mean of zero.

were dramatic in that 95% of the first grades and 100% of the third grades obtained the highest implementation score (see Table 9).

Table 9

CHILD GROUP RESPONSE TO ADULT ACADEMIC COMMANDS/REQUESTS OR DIRECT QUESTIONS* (Variable 363a)--UNIVERSITY OF OREGON

Sites	First Grade Classrooms with Implementation Scores of					Third Grade Classrooms with Implementation Scores of				
	1	2	3	4	5	1	2	3	4	5
E. St. Louis, Ill.					4					4
NYC P.S. 137K					3					3
Racine, Wis.			1		3					4
Tupelo, Miss.					4					4
Providence, R.I.					4					4
Total classrooms			1		18					18
Percent of classrooms			5%		95%					100%

*The sponsor refers to this as "Direct Instruction."

Adult praise was used systematically when the child response was correct. The child always knew immediately if his answer was right or wrong. Table 10 indicates that first grades had 63% and third grades had 79% of their classrooms in the fourth and fifth quintile on this variable. All of Racine's classrooms were in the fifth quintile. Only the classrooms in East St. Louis appeared atypical on this variable.

Table 10

ALL ADULT PRAISE TO CHILDREN (Variable 398a)--UNIVERSITY OF OREGON

Sites	First Grade Classrooms with Implementation Scores of					Third Grade Classrooms with Implementation Scores of				
	1	2	3	4	5	1	2	3	4	5
E. St. Louis, Ill.	2	1	1				2		2	
NYC P.S. 137K		2		1				1	1	1
Racine, Wis.					4					4
Tupelo, Miss.					4		1		1	2
Providence, R.I.			1	2	1				1	3
Total classrooms	2	3	2	3	9		3	1	5	10
Percent of classrooms	11%	16%	11%	16%	47%		16%	5%	26%	53%

Children in University of Oregon classrooms used workbooks and textbooks designed or recommended by the sponsor. As shown in Table 11, 79% of the first grades had a high implementation score of 4 or 5. While 47% of the third grades also had implementation scores of 5, the other classrooms were scattered across all of the scores. Only Tupelo had all four third grade classrooms with implementation scores of 5.

Table 11

TEXTS, WORKBOOKS/ACADEMIC ACTIVITIES (Variable 240)--UNIVERSITY OF OREGON

Sites	First Grade Classrooms with Implementation Scores of					Third Grade Classrooms with Implementation Scores of				
	1	2	3	4	5	1	2	3	4	5
E. St. Louis, Ill.				2	2	3				1
NYC P.S. 137K				1	2		1		1	1
Racine, Wis.		1	1	2			1	1	2	
Tupelo, Miss.			2	2						4
Providence, R.I.				1	3				1	3
Total classrooms		1	3	8	7	3	2	1	4	9
Percent of classrooms		5%	16%	42%	37%	16%	11%	5%	21%	47%

Summary of Implementation

The reader is reminded that a sponsor's implementation was evaluated in two ways: (1) Did they differ from the traditional classroom? and (2) Were a given sponsor's classrooms similar to each other in level of implementation? Using the first criterion our findings indicated that University of Oregon classrooms were statistically different from Non-Follow Through classrooms.

By the second criterion the first grades were implemented. The greatest deviation between sites was found in New York City (90%) and Racine (71%). East St. Louis, where one first grade had a score near to that of Non-Follow Through, had the greatest within-site variance. In third grade, while all classrooms but one had scores higher than Non-Follow Through classrooms, there was great deviation within and among sites. One explanation offered is that the teachers had altered the program in preparation for the next year in Non-Follow Through fourth grades.

Overall Summary

Our evaluation of the seven Follow Through sponsors in this study allowed no simple statement, such as, "All seven Follow Through sponsors succeeded in implementing their model in the sites which were studied." Rather, conditional statements were made. While all sponsor classroom mean scores for first and third grades were significantly different from the Non-Follow Through mean score, all sponsors except Far West Lab and University of Kansas and High/Scope had some classrooms which had scores below the Non-Follow Through mean. In both grades the University of Kansas implementation scores were higher than those of other sponsors.

The least amount of variance found among sites was in the Far West first grades (4.4). The most variance was found among the EDC first grade sites (11.5). Bank Street and High/Scope varied only moderately among sites in both grade levels.

Implementation of the models did not seem to be affected by the percent of children who did not use English as a first language; e.g., the scores in Salt Lake City, Fall River, Greeley, Denver, Fort Worth, and Lakewood were comparatively high.

Average low median family incomes did not seem to be related to successful implementation, e.g., Portageville, Rosebud, Greenwood, and Smithfield. All had low average incomes and high implementation scores.

The size of the city may have had some effect. Large cities accounted for seven of the eight implementation scores below 70 (see Table 12). The implementation scores for University of Oregon and High/Scope third grades in New York City were comparatively low and the between-classroom variance in New York was high for University of Oregon, High/Scope and University of Kansas. In Philadelphia, EDC had low implementation scores for both first and third grades as well as high within-site variance and Bank Street had low scores for three of the third grades. University of Kansas, however, had high implementation scores and little variation between classrooms in both grade levels in Philadelphia. Perhaps the teacher strikes in Philadelphia made it difficult to

Table 12

IMPLEMENTATION SCORES COMPARED BY SIZE
OF THE FOLLOW THROUGH SITES

Implementation Score	Rural	Town	Small City	Medium City	Big City	Total
90		2			1	3
85			1		4	5
80	1	4	4	4	2	15
75		1	5	9	9	24
70	1	1	4	6	5	17
65				1	4	5
60					2	2
55					1	1
50						
Total	2	8	14	20	28	72*

*Includes scores of both first and third grades.

implement the more open models represented by Bank Street and EDC since under stress there may be the tendency to become more structured. In Newark all of the first grades scored below the Non-Follow Through mean. The third grades scored close to, but not below, the Non-Follow Through mean. The variation between the classroom scores was very slight. The Newark site had all of the problems of the stereotypical inner-city school; the poverty rate was high, few adults over 25 had a high school education, the community varied in its opinion of the goals for the schools, and the sponsor reported difficulty in providing adequate classroom service during the first several years of sponsorship.

In general, small sites such as Smithfield, Greeley, Portageville, Tupelo, Tuskegee and Lebanon were implemented more consistently at both grade levels and they had higher implementation scores in both grades. This seems to indicate that the educational change represented by the Follow Through sponsors may be instituted more successfully in the smaller cities and towns than in the large inner-city schools. There are exceptions, of course; University of Kansas was successful in both grade levels in Philadelphia. Table 12 indicates that no rural area, town or small city had implementation scores lower than 70.

A Model for Program Implementation

This analysis of program implementation was accomplished in the early 1970s. The methodology developed over several years. The problem was to measure each model's classroom processes and determine whether teachers wer performing similar to each other within and among a sponsor's sites. The important contribution of the study was establishing a reliable yardstick by which each sponsor's classrooms could be measured.

TECHNICAL NOTE

Stability of Quintile Cutpoints

To assess the precision of the quintile estimates, 95 percent confidence intervals for the quintiles of the estimated classroom distribution were calculated. These are displayed for a typical variable in Table A.

Table A

95 PERCENT CONFIDENCE INTERVALS FOR THE QUINTILE ESTIMATES EXPRESSED IN PERCENTILES OF THE CLASSROOM DISTRIBUTION

Quintile Cutpoint	Corresponding Percentile	Confidence Interval
1	20	6 - 31
2	40	22 - 61
3	60	39 - 78
4	80	69 - 94

Consider the first quintile cutpoint as an example of how to interpret the confidence interval. If the experiment of sampling 36 classrooms over two or three days were replicated many times, then in 95 percent of the replications the first cutpoint, that is supposed to represent the 20th percentile, will be somewhere between the 6th and the 31st percentile of the Non-Follow Through classroom distribution.

These confidence intervals are rather wide, especially for the second and third quintiles. In terms of the implementation scores, these results indicate that for a given Follow Through classroom, the true implementation score may be plus or minus one unit from the observed score with a high degree of confidence. That is, if a classroom received a score of four, there is a good chance that the true score might be anywhere between three and five because of the variability of the quintile estimates. In interpreting these intervals it should be noted that observation on a single day might be less reliable (have wider confidence intervals). The effect of day to day variability in Non-Follow Through classrooms was examined and the intervals for the first two quintiles did shift downward, but the shift was only slight (see Table B). From these analyses we concluded that day to day variability has only a marginal effect on quintile estimates in comparison to the number of classrooms sampled. A more detailed explanation of these analyses is offered in Stallings and Kaskowitz, 1974.

Table B

THE VALUE OF l_o AND u_o EXPRESSED IN PERCENTILES
FOR EACH QUINTILE OF NON-FOLLOW THROUGH
(N = 36)

Quintile Cutpoint	Corresponding Percentile	Confidence Level $l_o \times 100$	$u_o \times 100$
1	20	10 -	34
2	40	26 -	59
3	60	41 -	74
4	80	66 -	90

REFERENCES

Charters, W. W. and Jones, J.E. On the risk of appraising non-events in program evaluation. Educational Researcher, 1973, 2 (11), 5-7

Egbert, R.L. Planned variation in Follow Through. Paper presented for the Brookings Institution Conference on Social Experimentation, Washington, D.C., April, 1973

Goodlad, J.C., Klein, M.F., et al. Behind the Classroom Door. Worthington, Ohio. Charles H. Jones Publishing Co., 1970

Mason, W. S. Problems of measurement and the NIE program. NIE mimeo, August, 1973

Stallings, J.A., Baker, P., and Steinmetz, G.T. Classroom observation study of implementation in Head Start planned variation, 1970-71. Stanford Research Institute, Menlo Park, California, 1973

Stallings, J.A. Follow Through Program Classroom Observation Evaluation. SRI Project URU-7370, Stanford Research Institute, Menlo Park, California, 1973

Stallings, J.A. and Kaskowitz, D. Follow Through Classroom Observation Evaluation 1972-1973. Stanford Research Institute, Menlo Park, California, 1974

Weikart, D.P. and Banet, B.A. Planned Variation: From the perspective of a model sponsor. Paper prepared for a working conference sponsored by the Brookings Institution Panel on Social Experimentation, Washington, D.C., April, 1973

Appendix A

Appendix A

UNIVERSITY OF OREGON : FREQUENCY DISTRIBUTION OF CLASSROOMS OVER IMPLEMENTATION SCORES

VARIABLE: CCL-A 46 NUMBERS, MATH, ARITHMETIC

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703				4				1		3
707				3			1	2		
708			2	2				2	2	
711			1	3						4
719		2			2		2	1		1
TOTAL		2		3	14		3	6	2	8
*PERCENT		11		16	74		16	32	11	42

VARIABLE: CCL-A A7 READING, ALPHABET, LANGUAGE DEVELOPMENT

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703		1	1	2			1	1	1	1
707				1	2	1			2	
708			4							4
711		1		3					1	3
719			4							4
TOTAL		2	9	3	5	1	1	1	4	12
*PERCENT		11	47	16	26	5	5	5	21	63

VARIABLE: CCL-B B8 TEACHER WITH SMALL GROUP

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703	1				3		1		1	2
707				1	2	1		1		1
708					4	2			1	1
711					4					4
719					4					4
TOTAL		1		1	17	3	1	1	2	12
*PERCENT		5		5	89	16	5	5	11	63

VARIABLE: CCL-B 96 AIDE WITH SMALL GROUP

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703		1	1	2						4
707			1	2				2	2	1
708				4			1	2	1	
711				4						4
719			1	3				1	3	
TOTAL			1	3	15		1	5	13	
*PERCENT			5	16	79		5	26	68	

VARIABLE: CCL-J 24A TEXTS, WOOKBOOKS/ ACADEMIC ACTIVITY

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703				2	2	3				1
707				1	2		1		1	1
708		1	1	2			1	1	2	
711			2	2						4
719				1	3				1	3
TOTAL		1	3	8	7	3	2	1	4	9
*PERCENT		5	16	42	37	16	11	5	21	47

VARIABLE: FHO-A 344 INDIVIUAL CHILD VERBAL INTERACTION/ ADULT

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703	4					4				
707		3				1	1		1	
708	3	1				1		1	2	
711	4					3		1		
719	4					3		1		
TOTAL	15	4				12	1	3	3	
*PERCENT	79	21				63	5	16	16	

VARIABLE: FHO-A 363 GROUP RESP. TO ADULT ACA. REW. OR DIR. QUES.

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703				4						4
707				3						3
708		1		3						4
711				4						4
719				4						4
TOTAL			1	18						19
*PERCENT			5	95						100

VARIABLE: FHO-A 376 ADULT INSTRUCTS GROUP

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703		4								4
707					3		3			
708	4						2	1		1
711	2		2						2	2
719		4				1	3			
TOTAL	6	8	2		3	1	8	1	2	7
*PERCENT	32	42	11		16	5	42	5	11	37

VARIABLE: FHO-A 394 ALL ADULT ACKN. TO CHILDREN

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703				2	2					4
707					3				1	2
708		2	2				3	1		
711		1	1		2		2		1	1
719	2	1	1				1	1	2	
TOTAL	2	4	4	2	7		6	1	1	4	7
*PERCENT	11	21	21	11	37		32	5	5	21	37

* PERCENT OF CLASSROOMS FOR A SPONSOR IN A GIVEN GRADE IN A GIVEN QUINTILE

UNIVERSITY OF OREGON 1 — FREQUENCY DISTRIBUTION OF CLASSROOMS OVER IMPLEMENTATION SCORES

VARIABLE: FMO-A 398 ALL ADULT PRAISE TO CHILDREN

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703	2	1	1				2		2	
707		2		1				1	1	1
708					4					4
711					4		1		1	2
719			1	2	1				1	3
TOTAL	2	3	2	3	9		3	1	5	10
*PERCENT	11	16	11	16	47		16	5	26	53

FMO-A 417 ADULT FDBK. TO CHILD RESP. TO ADULT ACA. CONN., QUES.

	FIRST					THIRD				
	1	2	3	4	5	1	2	3	4	5
				1	3				1	3
					3			1	1	1
					4				1	3
					4		1		1	2
	1	1		2						4
	1	1		3	14		1	1	4	13
	5	5		16	74		5	5	21	68

FMO-A 420 ADULT ATTEN. TO SMALL GROUP

	FIRST					THIRD				
	1	2	3	4	5	1	2	3	4	5
	1	1	1		1				1	3
					3				3	
					4		2			2
					4					4
			1		3					4
	1	1	2		15		2		4	13
	5	5	11		79		11		21	68

VARIABLE: FMO-A 435 TOTAL ACADEMIC VERBAL INTERACTIONS

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703				2	2			2		2
707				1	2		1	1	1	
708		4								4
711					4					4
719			1	1	2		1	1	1	
TOTAL		4	1	4	10		1	2	4	2
*PERCENT		21	5	21	53		5	11	21	11

FMO-A 44? ADULT FOCUS - SMALL GROUP

	FIRST					THIRD				
	1	2	3	4	5	1	2	3	4	5
	1			1	2				1	3
					3			1	1	
					4		1			3
					4					4
					4					4
	1			1	17		1	1	1	1
	5			5	89		5	5	5	5

FMO-A 451 ADULT ACA., CONN., REQ. ,DIR. TO CHILDREN

	FIRST					THIRD				
	1	2	3	4	5	1	2	3	4	5
					4		1	1	2	
					3					3
		1		3					1	3
					4					4
					4			1		3
		1		3	15		1	2	3	13
		5		16	79		5	11	16	68

VARIABLE: FMO-A 457 ALL ADULT POSITIVE CORRECTIVE FDBK.

GRADE:	FIRST					THIRD				
QUINTILE:	1	2	3	4	5	1	2	3	4	5
SITE										
703	2		2				4			
707			1	1	1		2			1
708	1	2	1				1	1		2
711		1		1	2		1	3		
719		2		2		4				
TOTAL	3	5	4	4	3		4	8	4	3
*PERCENT	16	26	21	21	16		21	42	21	16

FMO-C 509 CHILD SELF INSTRUCTION, ACADEMIC

	FIRST					THIRD				
	1	2	3	4	5	1	2	3	4	5
						1			2	1
							2		1	
									3	
						3	1			
						1			1	1
						5	3	1	7	2
						28	17	6	39	11

* PERCENT OF CLASSROOMS FOR A SPONSOR IN A GIVEN GRADE IN A GIVEN QUINTILE